



IFWO

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/723,148

DATE: 08/04/2004

TIME: 08:27:12

Input Set : N:\Crif3\RULE60\10723148.raw

Output Set: N:\CRF4\08042004\J723148.raw

1 <110> APPLICANT: Beraud, Christophe  
 2 Craven, Andrew  
 3 Yu, Ming  
 4 Sakowicz, Roman  
 5 Patel, Umesh A.  
 6 Davies, Katherine A.  
 7 <120> TITLE OF INVENTION: NOVEL MOTOR PROTEINS AND METHODS FOR THEIR USE  
 8 <130> FILE REFERENCE: 020552-001410US  
 9 <140> CURRENT APPLICATION NUMBER: US/10/723,148  
 10 <141> CURRENT FILING DATE: 2003-11-25  
 11 <150> PRIOR APPLICATION NUMBER: US/09/883,096  
 12 <151> PRIOR FILING DATE: 2001-06-15  
 13 <150> PRIOR APPLICATION NUMBER: US 09/594,655  
 14 <151> PRIOR FILING DATE: 2000-06-15  
 15 <160> NUMBER OF SEQ ID NOS: 6  
 16 <170> SOFTWARE: PatentIn Ver. 2.1  
 18 <210> SEQ ID NO: 1  
 19 <211> LENGTH: 4108  
 20 <212> TYPE: DNA  
 21 <213> ORGANISM: Artificial Sequence  
 22 <220> FEATURE:  
 23 <223> OTHER INFORMATION: Nucleic acid sequence of human kinesin motor  
 24 protein gene HsKip3a (Figure 1).  
 25 <223> OTHER INFORMATION: Description of Artificial Sequence: HsKip3a gene.

**ENTERED**

W--> 26 <400> 1  
 27 gggccgcga attcggcacc aggggcgctc tctcccggtg tgggtactgc tgtctgtggt 60  
 28 gtggctgtgg gaccctgtgag caagcagcga cgccagcggc ggagaaccga cgaaagggtgt 120  
 29 caccacagtg atggcagtgg aggacagcac gctgcaagta gtggtacggg tgcggccccc 180  
 30 caccctcgg gagctggaca gtcagcggcg gccagtgggt caggtggtgg acgagcgggt 240  
 31 gctggtgttt aaccctgagg agcccgatgg agggttccct ggctgaaat ggggtggcac 300  
 32 ccatgatggc cccaagaaga agggcaaaga cctgacgttt gtctttgacc gggctcttgg 360  
 33 cgaggcggcc acccaacagg acgtgttcca gcacaccacg cacagcgtcc tggacagctt 420  
 34 cctccagggc tacaactgct cagtgtttgc ctacggggcc accggggctg ggaagacaca 480  
 35 caccatgctg ggaagggagg gggaccccg catcatgtac ctgaccaccg tggaaactgta 540  
 36 caggcgcctg gaggcccgcc agcaggagaa gcacttcgag gtgctcatca gctaccagga 600  
 37 ggtgtataat gaacagatcc atgacctcct ggagcccaag gggcccttg ccatccgcga 660  
 38 ggaccccgac aagggggtgg tggtgcaagg actttcttcc caccagccag cctcagccga 720  
 39 gcagctgctg gagatactga ccagggggaa ccgtaaccgc acgcagcacc ccaactgatgc 780  
 40 caacgcgact tcctcccgct cccatgccat cttccagatc tttgtgaagc agcaggaccg 840  
 41 ggttccagga ctgaccagg ctgtccaggt ggccaagatg agcctgattg acctggctgg 900  
 42 ctcagagcgg gcatccagca cccatgcgaa gggggagcgg ctgcgggagg gggccaacat 960  
 43 caaccgctct ctgctggcgc tcatcaacgt cctcaatgcc ttggccgatg caaagggccg 1020  
 44 caagacccat gtgcctacc gggacagcaa actgaccgcg ctgctcaaag actccctcgg 1080

## RAW SEQUENCE LISTING

DATE: 08/04/2004

PATENT APPLICATION: US/10/723,148

TIME: 08:27:12

Input Set : N:\Crif3\RULE60\10723148.raw

Output Set: N:\CRF4\08042004\J723148.raw

```

45  gggcaactgc cgcacagtga tgatecgtgc catcagcccc tccagcctga cctacgagga 1140
46  cacgtacaac accctcaaat atgccgaccg ggccaaggag atcaggctct cgctgaagag 1200
47  caatgtgacc agcctggact gtcacatcag ccagtatgct accatctgcc aacagctcca 1260
48  ggctgaggtg gccgctctga ggaagaagct ccaagtgtat gaggggggag gccagcccc 1320
49  accacaggac ctcccaggat ctcccaagtc gggaccacca ccagaacacc ttcccagctc 1380
50  ccccttgcca cccaccctc ccagccagcc ctgcacccca gagctccctg cagggcctag 1440
51  agcccttcaa gaggagagtc tggggatgga ggcccagggtg gagagggcca tgggaaggaa 1500
52  ctcttcagac caggagcagt ccccagagga tgaggatgaa ggcccagctg aggaggttcc 1560
53  aaccagatg ccagagcaga accccacaca tgactgccca gagtccctc gcctgacct 1620
54  gcagcccaag ccagtcgtgg gccacttctc agcacgggaa ctggatgggg accgttctaa 1680
55  gcagttggcc ctaaagggtgc tgtgcgttgc ccagcggcag tactccctgc tccaagcagc 1740
56  caacctcctg acgcccagca tgatcacaga gtttgagacc ctacagcagc tggtgcaaga 1800
57  ggaaaaaatt gagcctgggg cagaggcctt gaggacttca ggccctggcca ggggggcacc 1860
58  tctggctcag gagctgtgtt cagagtcaat ccctgtgccg tctcctctct gccagagcc 1920
59  tccaggatac actggccctg tgaccgggac tatggcgagg cgactgagtg gccccctgca 1980
60  caccctggga atcccgcctg gacccaactg caccacagcc caggggtccc gatggcccat 2040
61  ggagaagaag aggaggagac caagcgcctt ggaggcagac agtcccatgg cctcaaagcg 2100
62  gggcaccaag cgccagcgcc agtccttctt gccctgccta aggagagggg ctctgcctga 2160
63  caccacaact tcacaggggc ccagcaccct caaaaggagaa agggcctcct ccccttgcca 2220
64  ttccctcctc gtttgcccag ccacagtcct caaaagccgg gtgcccctgg gcccttccgc 2280
65  catgcagaac tgctccacct cgctggctct gccactcga gacctcaatg ccacctttga 2340
66  tctctctgag gagcctccct caaagcccag ttccatgaa tgcattggct gggacaaaat 2400
67  accccaggag ctgagcaggc tggaccagcc cttcatcccc agggcacctg tgcccctgtt 2460
68  caccatgaag ggccccaaag caacatcttc cctccctggg acctctgcct gcaagaagaa 2520
69  gcgcgttgcg agttcctcag tctcccatgg ccgcagccgc atcgcccgcc tcccagcag 2580
70  cactttgaag aggccagctg ggccccttgt actccagag ctgcccttga gtcccctgtg 2640
71  ccctagcaac cggaggaatg gaaaggacct catcagggtg gggagagcgc tctcagcagg 2700
72  gaacggcgtc accaagggtg cctgaccgcc agaatgtcct gaccaccaag gtgtcctaac 2760
73  ctaccggccc ctctgctgga taccctctt ggacctgtag ccacctgcac caggagctgg 2820
74  acctgccttc cttacctggg agcaattagt gccaacacac ctttgctgta ttaacatccc 2880
75  tcccagaca tccatcctgc tactaccct ctgttaatct cctgttacac tcagcttctt 2940
76  ggcattgaca tattcatttg tgagtgttaa tgtgctgctg ttttttgttt tttggtggtt 3000
77  tttgtttttt gttttttttg ttttgagatg gagtcttact ctgtcgccca ggctggagt 3060
78  cagtggtagc atcttggtct actgcaacct ccgcctcctg ggttcaagta attctcctgc 3120
79  ctacgctttc caagtagctg ggattacagg caccatcac cacaccagc taattttcgt 3180
80  ctttttaata gagagggggg ttttccatgt tggccaggct ggtcttgaa tectgacctc 3240
81  aggtgatccg cctgcctcag cttcccaaag tgctgagatt acaggcatga gctaccacgc 3300
82  ctggcccggt ttgctgtttt aaagggtgct ccattgtccc ccatcttttt tttttttgag 3360
83  atggagtctc gctctgtcgc ccaggctgga gtgcagtggg ggcatcttg gctcactgca 3420
84  agtccgcct cccaggttca caccattctc ctgcctcagc ctcccaagta gctgggacta 3480
85  caggcgccca ccaccagcc cggctaattt tttgtatttt tagtagagat ggggtttcac 3540
86  cgtgttagcc aggtggtct cgtctgacc tcatgatcca ccgcctcgg cctcccaaag 3600
87  tgctgggatt acaggcgtga gccactgcgc ccggcctccc ctctcattta tgatgcctc 3660
88  tgtgcaggca gacggctctt gggctctttt cccacctgt ctctaacaca ggccccacgg 3720
89  tgatggccac aggcagtaga ggaggaatga ggatgggtg gggagcgggg agtcgcggct 3780
90  tggctcttcc tggtttctga gagggacatc ttcaccccta ctccccttgg tccccaacca 3840
91  cagtcctggg gaagatgtgg atgataatgg tgccttgatt tccaaatgaa gacagcttta 3900
92  ttgcttaact ctattgtaca taggatacac gttcagtgtg aaataaagtg taaaggggaa 3960
93  ttcaggctta atgctgcacc tagatataaa tgctaattgat acttgggttt atagccttct 4020

```

## RAW SEQUENCE LISTING

DATE: 08/04/2004

PATENT APPLICATION: US/10/723,148

TIME: 08:27:12

Input Set : N:\Crif3\RULE60\10723148.raw

Output Set: N:\CRF4\08042004\J723148.raw

```

94      gatcctttat ttctgcatat atatatagat atatacatat atttttggta taacaataaa 4080
95      ccgtctccat ccttgggaaa aaaaaaaaaa                                4108
97 <210> SEQ ID NO: 2
98 <211> LENGTH: 864
99 <212> TYPE: PRT
100 <213> ORGANISM: Artificial Sequence
101 <220> FEATURE:
102 <223> OTHER INFORMATION: Amino acid sequence encoded by human kinesin motor
103      protein gene HsKip3a (Figure 1).
104 <223> OTHER INFORMATION: Description of Artificial Sequence:Amino acid
105      sequence of HsKip3a.
W--> 106 <400> 2
107      Met Ala Val Glu Asp Ser Thr Leu Gln Val Val Val Arg Val Arg Pro
108           1             5             10             15
109      Pro Thr Pro Arg Glu Leu Asp Ser Gln Arg Arg Pro Val Val Gln Val
110           20             25             30
111      Val Asp Glu Arg Val Leu Val Phe Asn Pro Glu Glu Pro Asp Gly Gly
112           35             40             45
113      Phe Pro Gly Leu Lys Trp Gly Gly Thr His Asp Gly Pro Lys Lys Lys
114           50             55             60
115      Gly Lys Asp Leu Thr Phe Val Phe Asp Arg Val Phe Gly Glu Ala Ala
116           65             70             75             80
117      Thr Gln Gln Asp Val Phe Gln His Thr Thr His Ser Val Leu Asp Ser
118           85             90             95
119      Phe Leu Gln Gly Tyr Asn Cys Ser Val Phe Ala Tyr Gly Ala Thr Gly
120           100            105            110
121      Ala Gly Lys Thr His Thr Met Leu Gly Arg Glu Gly Asp Pro Gly Ile
122           115            120            125
123      Met Tyr Leu Thr Thr Val Glu Leu Tyr Arg Arg Leu Glu Ala Arg Gln
124           130            135            140
125      Gln Glu Lys His Phe Glu Val Leu Ile Ser Tyr Gln Glu Val Tyr Asn
126           145            150            155            160
127      Glu Gln Ile His Asp Leu Leu Glu Pro Lys Gly Pro Leu Ala Ile Arg
128           165            170            175
129      Glu Asp Pro Asp Lys Gly Val Val Val Gln Gly Leu Ser Phe His Gln
130           180            185            190
131      Pro Ala Ser Ala Glu Gln Leu Leu Glu Ile Leu Thr Arg Gly Asn Arg
132           195            200            205
133      Asn Arg Thr Gln His Pro Thr Asp Ala Asn Ala Thr Ser Ser Arg Ser
134           210            215            220
135      His Ala Ile Phe Gln Ile Phe Val Lys Gln Gln Asp Arg Val Pro Gly
136           225            230            235            240
137      Leu Thr Gln Ala Val Gln Val Ala Lys Met Ser Leu Ile Asp Leu Ala
138           245            250            255
139      Gly Ser Glu Arg Ala Ser Ser Thr His Ala Lys Gly Glu Arg Leu Arg
140           260            265            270
141      Glu Gly Ala Asn Ile Asn Arg Ser Leu Leu Ala Leu Ile Asn Val Leu
142           275            280            285
143      Asn Ala Leu Ala Asp Ala Lys Gly Arg Lys Thr His Val Pro Tyr Arg

```

## RAW SEQUENCE LISTING

DATE: 08/04/2004

PATENT APPLICATION: US/10/723,148

TIME: 08:27:12

Input Set : N:\Crif3\RULE60\10723148.raw

Output Set: N:\CRF4\08042004\J723148.raw

144		290		295		300													
145	Asp	Ser	Lys	Leu	Thr	Arg	Leu	Leu	Lys	Asp	Ser	Leu	Gly	Gly	Asn	Cys			
146	305					310					315					320			
147	Arg	Thr	Val	Met	Ile	Ala	Ala	Ile	Ser	Pro	Ser	Ser	Leu	Thr	Tyr	Glu			
148					325					330						335			
149	Asp	Thr	Tyr	Asn	Thr	Leu	Lys	Tyr	Ala	Asp	Arg	Ala	Lys	Glu	Ile	Arg			
150				340					345					350					
151	Leu	Ser	Leu	Lys	Ser	Asn	Val	Thr	Ser	Leu	Asp	Cys	His	Ile	Ser	Gln			
152			355					360					365						
153	Tyr	Ala	Thr	Ile	Cys	Gln	Gln	Leu	Gln	Ala	Glu	Val	Ala	Ala	Leu	Arg			
154		370					375				380								
155	Lys	Lys	Leu	Gln	Val	Tyr	Glu	Gly	Gly	Gly	Gln	Pro	Pro	Pro	Gln	Asp			
156	385					390				395						400			
157	Leu	Pro	Gly	Ser	Pro	Lys	Ser	Gly	Pro	Pro	Pro	Glu	His	Leu	Pro	Ser			
158					405					410						415			
159	Ser	Pro	Leu	Pro	Pro	His	Pro	Pro	Ser	Gln	Pro	Cys	Thr	Pro	Glu	Leu			
160				420					425					430					
161	Pro	Ala	Gly	Pro	Arg	Ala	Leu	Gln	Glu	Glu	Ser	Leu	Gly	Met	Glu	Ala			
162			435					440					445						
163	Gln	Val	Glu	Arg	Ala	Met	Glu	Gly	Asn	Ser	Ser	Asp	Gln	Glu	Gln	Ser			
164		450					455					460							
165	Pro	Glu	Asp	Glu	Asp	Glu	Gly	Pro	Ala	Glu	Glu	Val	Pro	Thr	Gln	Met			
166	465					470				475						480			
167	Pro	Glu	Gln	Asn	Pro	Thr	His	Ala	Leu	Pro	Glu	Ser	Pro	Arg	Leu	Thr			
168				485						490						495			
169	Leu	Gln	Pro	Lys	Pro	Val	Val	Gly	His	Phe	Ser	Ala	Arg	Glu	Leu	Asp			
170			500						505					510					
171	Gly	Asp	Arg	Ser	Lys	Gln	Leu	Ala	Leu	Lys	Val	Leu	Cys	Val	Ala	Gln			
172			515					520					525						
173	Arg	Gln	Tyr	Ser	Leu	Leu	Gln	Ala	Ala	Asn	Leu	Leu	Thr	Pro	Asp	Met			
174		530					535					540							
175	Ile	Thr	Glu	Phe	Glu	Thr	Leu	Gln	Gln	Leu	Val	Gln	Glu	Glu	Lys	Ile			
176	545					550				555						560			
177	Glu	Pro	Gly	Ala	Glu	Ala	Leu	Arg	Thr	Ser	Gly	Leu	Ala	Arg	Gly	Ala			
178				565						570						575			
179	Pro	Leu	Ala	Gln	Glu	Leu	Cys	Ser	Glu	Ser	Ile	Pro	Val	Pro	Ser	Pro			
180			580						585					590					
181	Leu	Cys	Pro	Glu	Pro	Pro	Gly	Tyr	Thr	Gly	Pro	Val	Thr	Arg	Thr	Met			
182			595					600					605						
183	Ala	Arg	Arg	Leu	Ser	Gly	Pro	Leu	His	Thr	Leu	Gly	Ile	Pro	Pro	Gly			
184		610					615					620							
185	Pro	Asn	Cys	Thr	Pro	Ala	Gln	Gly	Ser	Arg	Trp	Pro	Met	Glu	Lys	Lys			
186	625					630				635						640			
187	Arg	Arg	Arg	Pro	Ser	Ala	Leu	Glu	Ala	Asp	Ser	Pro	Met	Ala	Ser	Lys			
188				645						650						655			
189	Arg	Gly	Thr	Lys	Arg	Gln	Arg	Gln	Ser	Phe	Leu	Pro	Cys	Leu	Arg	Arg			
190			660						665					670					
191	Gly	Ser	Leu	Pro	Asp	Thr	Gln	Pro	Ser	Gln	Gly	Pro	Ser	Thr	Pro	Lys			
192			675					680					685						

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/723,148

DATE: 08/04/2004

TIME: 08:27:12

Input Set : N:\Crif3\RULE60\10723148.raw

Output Set: N:\CRF4\08042004\J723148.raw

```

193   Gly Glu Arg Ala Ser Ser Pro Cys His Ser Pro Arg Val Cys Pro Ala
194       690               695               700
195   Thr Val Ile Lys Ser Arg Val Pro Leu Gly Pro Ser Ala Met Gln Asn
196   705               710               715               720
197   Cys Ser Thr Pro Leu Ala Leu Pro Thr Arg Asp Leu Asn Ala Thr Phe
198               725               730               735
199   Asp Leu Ser Glu Glu Pro Pro Ser Lys Pro Ser Phe His Glu Cys Ile
200               740               745               750
201   Gly Trp Asp Lys Ile Pro Gln Glu Leu Ser Arg Leu Asp Gln Pro Phe
202       755               760               765
203   Ile Pro Arg Ala Pro Val Pro Leu Phe Thr Met Lys Gly Pro Lys Pro
204       770               775               780
205   Thr Ser Ser Leu Pro Gly Thr Ser Ala Cys Lys Lys Lys Arg Val Ala
206   785               790               795               800
207   Ser Ser Ser Val Ser His Gly Arg Ser Arg Ile Ala Arg Leu Pro Ser
208               805               810               815
209   Ser Thr Leu Lys Arg Pro Ala Gly Pro Leu Val Leu Pro Glu Leu Pro
210               820               825               830
211   Leu Ser Pro Leu Cys Pro Ser Asn Arg Arg Asn Gly Lys Asp Leu Ile
212       835               840               845
213   Arg Val Gly Arg Ala Leu Ser Ala Gly Asn Gly Val Thr Lys Val Ser
214       850               855               860

```

216 &lt;210&gt; SEQ ID NO: 3

217 &lt;211&gt; LENGTH: 1014

218 &lt;212&gt; TYPE: DNA

219 &lt;213&gt; ORGANISM: Artificial Sequence

220 &lt;220&gt; FEATURE:

221 &lt;223&gt; OTHER INFORMATION: Description of Artificial Sequence: HsKip3a

222 fragment

223 &lt;223&gt; OTHER INFORMATION: Nucleotide sequence encoding motor domain fragment

224 of HsKip3a (Figure 2).

W--&gt; 225 &lt;400&gt; 3

```

226   gacagcacgc tgcaagtagt ggtacgggtg cggcccccca cccctcggga gctggacagt 60
227   cagcggcggc cagtggttca ggtggtggac gagcgggtgc tgggtgttaa ccctgaggag 120
228   cccgatggag ggttccctgg cctgaaatgg ggtggcaccc atgatggccc caagaagaag 180
229   ggcaaagacc tgacgtttgt ctttgaccgg gtctttggcg aggcggccac ccaacaggac 240
230   gtgttccagc acaccacgca cagcgtctg gacagcttcc tccagggcta caactgctca 300
231   gtgtttgcct acggggccac cggggctggg aagacacaca ccatgctggg aaggaggagg 360
232   gaccccgga tcatgtacct gaccaccgtg gaactgtaca ggcgcctgga ggcccgcag 420
233   caggagaagc acttcgaggt gctcatcagc taccaggagg tgtataatga acagatccat 480
234   gacctcctgg agcccaagg gccccttgcc atccgcgagg accccgacaa ggggggtggtg 540
235   gtgcaaggac tttctttcca ccagccagcc tcagccgagc agctgctgga gatactgacc 600
236   agggggaacc gtaaccgcac gcagcacccc actgatgcca acgcgacttc ctcccgtctc 660
237   catgccatct tccagatctt tgtgaagcag caggaccggg ttccaggact gaccaggct 720
238   gtccaggtgg ccaagatgag cctgattgac ctggctggct cagagcgggc atccagcacc 780
239   catgcaagg gggagcggct gcgggagggg gccaacatca accgctctct gctggcgctc 840
240   atcaacgtcc tcaatgcctt ggccgatgca aagggccgca agacccatgt gccctaccgg 900
241   gacagcaaac tgaccgcct gctcaaagac tccctcgggg gcaactgccg cacagtgatg 960
242   atcgctgcca tcagcccctc cagcctgacc tacgaggaca cgtacaacac cctc 1014

```

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/723,148

DATE: 08/04/2004

TIME: 08:27:13

Input Set : N:\Crf3\RULE60\10723148.raw

Output Set: N:\CRF4\08042004\J723148.raw

L:26 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:1  
L:106 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:2  
L:225 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:3  
L:252 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:4  
L:306 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:5  
L:365 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:6